

Copyright (c) 1993 - 2000 Compugen Ltd.  
GenCore version 4.5

## OM protein - protein search, using sw model

Run on: September 11, 2002, 12:13:28 ; Search time 39.17 Seconds  
(without alignments)  
2852,639 Million cell updates/sec

Title: US-09-841-132-190  
Perfect score: 5238  
Sequence: 1 MASMTGGQQMGRDSSLVPHH.....YERDASRGYGLSAGSKVRF 1006

Scoring table: BLOSUM62  
Gapext 0.5  
Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 95%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802:  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:/\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:/\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:/\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:/\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:/\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:/\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:/\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:/\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:/\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:/\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:/\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:/\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:/\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:/\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:/\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:/\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:/\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:/\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:/\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:/\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:/\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5238	100.0	1006	22 AAG83207	Protein encoded by C. trachomatis pmp
2	5231	99.9	1006	21 AAB13639	C. trachomatis pmp
3	5094	97.3	982	21 AAB13633	Protein encoded by C. trachomatis LGV
4	5094	97.3	982	22 AAG83201	
5	5090	97.2	1012	20 AAY16735	

## ALIGNMENTS

QY	1 MASMTGGQQMGRDSSLVPHH.....YERDASRGYGLSAGSKVRF 1006	Query Match 100.0% ; Score 5238; DB 22; Length 1005;
QY	1 masmtggggmgrosslyvhphhhmipgiydetitfsptvlgdsgstlvfsaglt 60	Best Local Similarity 100.0% ; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 1006; Conservative 0;
Db	1 masmtggggmgrosslyvhphhhmipgiydetitfsptvlgdsgstlvfsaglt 60	
QY	1 LKNLDNSTAALPLSCFGNLLGSPFLVIGRHSLTTFENITSTNGAALSNSAADCGLFTIEGF 120	
Db	61 1knldnsaalplscfgnllgspflvigrhslttfenitstngaaalsnsaadcglftiegf 120	
QY	1 KELFSFNSNCSLLAVLPATTNKSQTPTTSTNSGTYSKTDLILLNEKEFSEYSNLLVS 180	
Db	121 kelsfsnnsllavlpattnksqtpttstnsngtysktdlillnekefsysnlvs 180	
QY	181 GDGGAIDAKSLTIVQGISKLCVFOEINTAAQDGACQWVTSFSANANEAPTAIVANVAGYRG 240	
Db	181 gdggaidsksltivqgkislkcvfqentaaqdgacqvvfsananeaptaivavagrg 240	
QY	241 GGIAAVQDGQQGQYSSSTTEDPVVSFSRNTAVEDGVNARVGGIYSYGNVAFINNGKTL 300	
Db	241 ggiaavqdgqqgqysssttedpvvsfsrntaveldgvnrvarggqiyssygnvafinngktl 300	

QY	301	FLNNVAPVYIAAKOPTSQASNTSNNYGDGAIFCKNGAOGAQSNNSGSVSTFDGEVYVFF	360	PR	08-DEC-1998;	98US-0208277.
Db	301	flnnvapvyaakoptsqasntsnnygdaifckngaoqasnnsgsvstfdgevffff	360	PR	08-APR-1999;	99US-0208594.
QY	361	SNYAAGKGGAIAKKSYANCGPVQFLRNIAANDGATYLGESEGELSUSADGDIIFDGN	420	PR	01-OCT-1999;	99US-0140568.
Db	361	snvaagkggaiyakksvancpvgfirlnianandggaiygegelsisadgdiifdg	420	PR	22-OCT-1999;	99US-0426571.
QY	421	LKRTAKENADVNQGTVSSQATNSGGKITDRAKASHOILENDPIEMANGNQPAQSS	480	XX	(CORI- ) CORIYA CORP.	
Db	421	lkrtakenaadvnqgtvssqatnsggkitdrakashoilendpiemangnqpaqss	480	XX	Probst, P., Bhatia, A., Skeiky, Yaw, Fling, SP., Jen, S., Stromberg, EJ;	
QY	481	KLLINDGESEYGDIVFANGSTLQNTYIEQGRIVLREKAKLVSNSLQTEGSLYMEAG	540	XX	WPI; 2000-431303/37.	
Db	481	kllindgegygdifangstlyqntyieqgrivlrekaklvsnslqtegslymeag	540	XX	Isolated polypeptide for diagnosis and treatment of Chlamydia infection	
QY	541	STUDFVTPOPPOOPPAANOLITLISNLHLSLSLIANNAVTNPPPTNPAQDSDHPIAVGISTT	600	XX	comprises immunogenic portion of Chlamydia antigen, which comprises	
Db	541	studfvtpoppoppaanolitlislhlsliannavtnppptnppaqdshpavgit	600	XX	amino acid sequence encoded by polynucleotide sequence	
QY	601	AGSYTISGPFPEELDDRAYDRLWLSNOKINVKLQLGTRPANAPSDLTGEMNPKY	660	CC	Claim 2; Pages 208-210; 256pp; English.	
Db	601	agsvlisqpfedlddaydrdlwlsnokinvklqlgtpranapsdltgemnpky	660	CC	the present invention relates to new nucleic acid sequences and the	
QY	661	GYQSWKFLAWDPNTANNGPYTLKATWTKGYNNGPERVASYLNSLNGSILDILRSAAI	720	CC	proteins encoded by the nucleic acid sequences. The encoded proteins	
Db	661	gyqswkflawdpntanngpytlkawtktgyngperaslrvpslsgsildlrsaaai	720	CC	comprise an immunogenic portion of a Chlamydia antigen. The encoded	
QY	721	QASVGDGRSYCRGLWVGSYNSNFYHDRDALQGQYRISGGYSLGANSYFGSSMFLGAFTEV	780	CC	proteins are useful for the serodiagnosis and treatment of Chlamydia	
Db	721	qasvdgrsycrglwvgsynsfyhdralqgyryisggysgansyegsnglaftev	780	CC	infection. Chlamidae are intracellular bacterial pathogens that are	
QY	781	FGRSKDYYCRSNHACIGSVYLSTQQLCGSYLFGDAFTRASYGFQHNMHTSYFAE	840	CC	responsible for a wide variety of human infections. C. trachomatis	
Db	781	fgrskdyyvcrsnhacigsvylstqqlcgsvyfqaafirasyegnqhmhtsyfae	840	CC	infection is one of the most common sexually transmitted diseases and can	
QY	841	SDVRWDNNCLAGEGAGLPIVITPSKLYNLERPFPVQAEFSTADHESSTEEDQARAFKS	900	CC	lead to pelvic inflammatory disease (PID), resulting in tubal obstruction	
Db	841	sdvrdwnnclagegaglpivitpsklynlernfpvqaefstadhesssteedqarafks	900	CC	and infertility. Trachoma due to ocular infection with C. trachomatis	
QY	901	GHILNLNSYEPVGKEDRCSSSTHPNKSYNSAAYCDAYRTISGTETTLISHOETWTDASHL	960	CC	is the leading cause of preventable blindness worldwide. C. pneumonia is a	
Db	901	ghilnlnsyepvgkfdrcssthpnksyasnaycdayrtisgtettlishoetwdashl	960	CC	major cause of acute respiratory tract infections in humans and is also	
QY	961	ARAGYYVRSYMSYASLTSNIEVYHGRYRYRDAARGYGIISAGSYVRF	1006	CC	thought to play a role in the pathogenesis of atherosclerosis and	
Db	961	arhgvvrgsmyasltsievyhgryrydarsrgyisagskvr	1006	CC	coronary heart disease. The present sequence is a protein isolated in the	
Db			CC	present invention.		
QY			CC	Sequence 1006 AA:		
Db			CC	Query Match 99.9%; Score 5231; DB 21; Length 1006;		
QY			CC	Best Local Similarity 99.9%; Pred. No. 0;		
Db			CC	Matches 1005; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY			QY	1 MASMTGGOMGRDSSLYFHMHMIGQIYDGETLVSFPTVIGPSGTVFSEAGELT 60		
Db			Db	1 masmtggomgrdsslyfhmhmigqiydgetlvsfptvigpsgtlvfseagel 60		
QY			QY	61 LKNDNSAALPLSFCFGNLLGSFTVLGRHSLSFPENIRTSNGAALSNSAAGLFTBGF 120		
Db			Db	61 lkndnsaalplscfcgnllgsftvlgrhslsfpenirtsngalsnsaaglf 120		
QY			QY	121 KELSFNSNCNSLIAVLPAATTNGKSGQTPTTSTPSNGTNTYSKTDLLINNEKEFSYNSLVS 180		
Db			Db	121 kelsfnsncnsliavlpattngksqtpbttspsngtysktdlllnmekfsfyslvs 180		
QY			QY	181 GDGAIDAKSLTIVQGISKLCYFOENTAQADGEGACQVYTSFSAMANEAPIAFYANVAVGVRG 240		
Db			Db	181 gdgaidaksltivqgisklcycqntafvqntarefdgvnvarqgyisqnvafnlngktl 300		
QY			QY	241 GGAAVQDGQGQYSSSSTEDPVFSRNTAVFEDGVNARYGGITSYGNVAFLNNGKTL 300		
Db			Db	241 ggaaavqdgqgysssstedpvfsrntavfedgvnaryggisqnvafnlngktl 300		
QY			QY	301 FLNNVASVYIAKOPTSGQASNTSNNYGDGAIFCKNGAQAGSNNNSGSVSPDGEVYFFF 360		
Db			Db	301 flnnvasvyyiaakoptsgqasntsnnygdaifckngaqagsnssvsfdgegvff 360		
QY			QY	361 SNNYAAIGGGAYAKKSLVANCGPYQFLRNIAANDGATYLGESELSADYGDILFDGN 420		
Db			Db	361 ssivaaggayaiaksvancpvgfirlnianandggaiygegelsisadgdiifdrgn 420		
QY			QY	421 LKRTAKENADVNQGTVSSQASMGSGKKITLRAKAGHQILFNDPLIEMANGNNQPAQSS 480		
Db			Db	421 lkrtakenaadvnqgtvssqalsmgsqkittlrakahqilfndpliemangnnqpaqs 480		
QY			QY	481 KLLKINDGEYGDIVFANGSTLYQNTIEGRVILREKAKLVSNSLSPQGSLYMEAG 540		
XX			XX			

Db 481 killkindgegytqdivfangsstlyqnvttiegrivlrekaklsvnslsqtgtsglymeag 540  
 QY 541 STLDFVTPQPPOPPAANOLITLNSLNHLSSLANNAVNTNPPNPAODSHPAVIGSTT 600  
 Db 541 stldfvtpqppoppaanolitlnslnhsslanavntnppnqaodspavigstt 600  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 QY 601 AGSYTISGPIFEDDDTAYDRDWLGSNOKINVKQLQLGTPKANPAPSDLTIGNEMPKY 660  
 Db 601 agsytisgpifedddtaydrdwlgnsqknlqgktpkanspdsdlqnmeky 660  
 QY 661 GYOGSWKLAQDPNTANNGPYTAKATWTKTGYNPGPERVASLVPNSLMGSILDIRSAHSI 720  
 Db 661 gyogswklaqdpntanngpytakatwtktgynpgpervaslvpnsiwdqslidirsaisi 720  
 QY 721 QASVYDGRSYCQGLWVSGYSNFYHQRDHALGOQYRISGGYSLGANSYFQGSSMFGLAFTEV 780  
 Db 721 qasvdgrsyrcgqglwvsgysnfyhydrdhalgoqyrisggysgansyfsgmifglatev 780  
 QY 781 FGRSKDYYVCRSNHHACTGSVYLSTQOALCGSYLFGDAFIRASYGFENQHMTSYTFAEE 840  
 Db 781 fgrskdyyvcrsnhhactgsvylstqoalcgsylyfgdafirasygfonqhmitsytaee 840  
 QY 841 SDYRWDNNCLAGEIGAGLPIVITPSKLYNLERPFPVQAEFSSADHESFTEEDQARAFK 900  
 Db 841 sdyrwdnnclageigaglpivitpsklynlrpfvqaefsyaadhesfteeedqarafk 900  
 QY 901 GHLLNLSVPVGKRFDRCSSTTHPKYFSMAAYTCDAYRTISGTETLISHQEWTDAEHL 960  
 Db 901 ghllnlsvpvgkrfdrccsstthpkysfmaaytcdayrtisgtetlshqewttdfhl 960  
 QY 961 ARHGVVVYRGSMYASLTSNTSNEVYKGRYEYRDAASRGYGLSAGGSKVR 1006  
 Db 961 arhgvvvrgsmyslttsnevykgryeysrdsrgyglsgasgvkvr 1006  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 AC AAB13633 standard; Protein; 982 AA.  
 XX  
 AC AAB13633;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE C. trachomatis pmpG gene protein.  
 XX  
 DE Chlamydial infection; sexually transmitted disease;  
 KW pelvic inflammatory disease; PID; tubal obstruction; infertility;  
 KW trachoma; blindness; acute respiratory tract infection;  
 KW atherosclerosis; coronary heart disease; antibacterial.  
 XX  
 OS Chlamydia trachomatis.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 981  
 FT /note= "Unspecified amino acid"  
 XX  
 PN WO200034483-A2.  
 XX  
 PD 15-JUN-2000.  
 XX  
 PR 08-DEC-1999; 99WO-US29012.  
 XX  
 PR 08-DEC-1998; 98US-0268277.  
 PR 08-APR-1999; 99US-028594.  
 PR 01-OCT-1999; 99US-0410568.  
 PR 22-OCT-1999; 99US-0426571.  
 XX  
 PA (CORT-) CORIXA CORP.  
 XX  
 PI Probst P, Bhatia A, Skeiky YAW, Fling SP, Jen S, Stromberg EU,  
 XX  
 DR W15N9QKINVKQVLQGTPKANPAPSDLTIGNEMPKYQGWSNLAWDENTANNGPYTIKA 684  
 XX  
 DR 601 w15n9qkinvkqvlqgtpkanspds1tgnempkyqgqswlawdntanngpytika 660  
 XX  
 DR 685 TWTKTGYNGPGPERVASLVPNSLQGSLIDIRSAHSIAQASVDDGSYCRGLWVSGVSNFYH 744

PT Isolated polypeptide for diagnosis and treatment of Chlamydia infection  
 PT comprises immunogenic portion of Chlamydia antigen, which comprises  
 PT amino acid sequence encoded by polynucleotide sequence -  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksivvqgklycivqf 180  
 QY 205 NTQAQDGACQVNTFSAMANAEPIAFVANAVGVRGGGIAAYODGQGQVSSSTSSTEDPPV 264  
 Db 181 ntqaqdgacqvtfsamaneapiatvavqgrgq9iaarvqdgqgq9vssstsedppv 240  
 QY 265 SFERNATAEFDGNAVARVGGIYSGVAFLNNGKTLFLINNVASPVYIAKQPTSGQASNT 324  
 Db 241 sfsrntataefdgnavarvggqiygsvflngktlflinnavspvyiaakqptsgqasnt 300  
 QY 325 SNYYGDGAIIFCKNGAQAGSNNSGSVSYSDGEVYFFSSVNAACKGGAYAKLISVANGCP 384  
 Db 301 snyygdgaiifckngaqgsnnnsvsfdgegrvffsnnvaqgkgyaklislsvangcp 360  
 QY 385 VQFLRNIAANDGGAYIYGEGELSADYDGFNLKRTAKENAADVNGTVSSQALSM 444  
 Db 361 vqflrniandgayiygsgelsadsydgfndqnlkrtakenaadvngtvssqalsm 420  
 QY 445 GSGGKRTITLRAKQHQLIENDPLEMANGNNQPAQSSKLLKINDGEVGTGDIVPANGSTL 504  
 Db 421 gsgkrtitlrakeghqilfindpiemangnnqpaqsskllkindgeygtdivfangstl 480  
 XX  
 PS Claim 2; Pages 181-184; 256pp; English.  
 XX  
 CC The present invention relates to new nucleic acid sequences and the  
 CC proteins encoded by the nucleic acid sequences. The encoded proteins  
 CC comprise an immunogenic portion of a Chlamydia antigen. The encoded  
 CC proteins are useful for the serodiagnosis and treatment of Chlamydia  
 CC infection. Chlamydia are intracellular bacterial pathogens that are  
 CC responsible for a wide variety of human infections. C. trachomatis  
 CC infection is one of the most common sexually transmitted diseases and can  
 CC lead to pelvic inflammatory disease (PID), resulting in tubal obstruction  
 CC and infertility. Trachoma due to ocular infection with C. trachomatis is  
 CC the leading cause of preventable blindness worldwide. C. pneumonia is a  
 CC major cause of acute respiratory tract infections in humans and is also  
 CC thought to play a role in the pathogenesis of atherosclerosis and  
 CC coronary heart disease. The present sequence is a protein isolated in the  
 CC present invention.  
 XX  
 SQ Sequence 982 AA;  
 SQ  
 Query Match 97.3%; Score 5094; DB 21; Length 982;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 25 MIPOGIVNGETLTVSFPTVIGDPSGMTVFSAGETLTKNLNDNSIAALPSCFGNLGSFT 84  
 Db 1 mlpqgivngetltvsfptvlgdpsgtvfgseltlknlnsdiaalpclsqgnlgsft 60  
 QY 85 VLGRGHSLTIFENIRTSNGAALSNSAAGDGLFTEGKELSEFSNCNSCLAVLPATTNGS 144  
 Db 61 vlgrhsltfenvirtsngaaalsnsaagdglfiegtkelsfsncnsllavlpattngs 120  
 QY 145 QPRTTTSPSNGTIYSKIDLILLNNEKEFSYSNLVSQDGGIDAQKSLTVOGTSKLCYQE 204  
 Db 121 qptttspsngtlyskidlillnnkefsylnlvsqggaidsksiv

Db	661	twtktgyngppervaslypnslwsldirsahsaijasvdrgrscrglwvgsnifyh	720	Query Match 97.3%; Score 5094; DB 22; Length 982;
QY	745	DRLAQLQGYRYISCGYSUANCYFGSSMFLGALATEVERSKIDVVCRSNHHACIGSYLS	804	Best Local Similarity 99.9%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Db	721	drdalqgyryisggysgansfgrsmflgatlevfrsksdvvcrsnhhacigsylys	780	Matches 981; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	805	TQOALCGSYLFLGDAFIRASYFGFQNHMKTSYTAESDVRWNCLAGEIGAGLPIVITP	864	QY 25 MFPQGITYGETLTVSEPTVVGDSGTTVFSAGELTLURNLNDNSTIAALPLSCFGNLQGST 84
Db	781	tqqlcgsylfgafairsgyfgrqmktsytaeesvrvdnncleigggipivtp	840	Db 1 mppgiygettvsfpvtvqdpstvfgseltlqnlndnsiaalplsfcgnlqfst 60
QY	865	SKLYLNELRPFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSVPGVFKEDRSSTHPK	924	QY 85 VLRGRHSLTFENIRTSNGAALNSAADCGLTIEGFKELSFNSNCNSLAVLPATTNGS 144
Db	841	sklylnelrpfvqaefsyadhesiteedqarafksghllnlsvpgvfkdrccssthpnk	900	Db 61 vgrghstfenvirtsngalnsaadcgltiegfkelsfnccnslaviapattngs 120
QY	925	YFMAAYICDAYRTISGTTTSHQETWTTDAFHLLRHGVVVRGSMYASLSNIEVYGH	984	QY 145 QTPTTTSPSNTIYSKTDJLILNNEKSFYSMLVSDGGAQDAKSUTVQGISKICVYQE 204
Db	901	yfmaayicdayrtisgtttshqetwttdafhllrhgvvvrsgmyasitsnievgh	960	Db 121 qptttspsngtiysskdlilnneksfyslvsndggaidaksitvqgkislcvfcqe 180
QY	985	GYEYRDAASRGYGLISAGSKYRF	1006	QY 205 NTQADGACQVVTFSMSMANEAPIAFVANYAVRGCCGIAAVQDGQGVSSSTSDEDPPV 264
Db	961	gyeyrdaasrgyglisagskysvxf	982	Db 181 ntqadgacqvvtsfamaneapiavanyavrg99iaavqdgq99vssststdepvv 240
QY	265	SFERNTAYEFDGNVARYGGIYSYGNVAFLNNGKTLFLNNVASPVITAKOPTSGQASNT	324	QY 265 SFERNTAYEFDGNVARYGGIYSYGNVAFLNNGKTLFLNNVASPVITAKOPTSGQASNT 324
Db	241	sfsrntavefdgivavrgginsygnvaflmngktflnnvaspvyiakoptsgqasnt	300.	Db 241 sfsrntavefdgivavrgginsygnvaflmngktflnnvaspvyiakoptsgqasnt 300.
QY	325	SNNGDGGATFCKNGAQAGSNNSGSVSFDGEVVFESSVNAAGKGGAAYAKLSVANC GP	384	QY 325 SNNGDGGATFCKNGAQAGSNNSGSVSFDGEVVFESSVNAAGKGGAAYAKLSVANC GP 384
Db	301	snnygdgafckngaqgsnsqsvsdgegrvftssnvaqgkgaayaklsvanc gp	360	Db 301 snnygdgafckngaqgsnsqsvsdgegrvftssnvaqgkgaayaklsvanc gp 360
QY	385	VOPLRNIAANDGGAIALGSGEELSADYGDIDFDGNLURKRTAKENADYNGVTVSSQAIM	444	QY 385 VOPLRNIAANDGGAIALGSGEELSADYGDIDFDGNLURKRTAKENADYNGVTVSSQAIM 444
Db	361	vqfrrnlandggaaylyqsgelslsadygqsfqnlkrtakenadnygvtssqaim	420	Db 361 vqfrrnlandggaaylyqsgelslsadygqsfqnlkrtakenadnygvtssqaim 420
QY	445	GSGKIRMLRAKAGHQTFLNDPLEMANGNQPAQSSKLLKINGEGTYDVFYANGSTL	504	QY 445 GSGKIRMLRAKAGHQTFLNDPLEMANGNQPAQSSKLLKINGEGTYDVFYANGSTL 504
Db	421	gsgkirktrakkhqlfnpliemangnqpaqsskllkingegtydvfangstl	480	Db 421 gsgkirktrakkhqlfnpliemangnqpaqsskllkingegtydvfangstl 480
QY	505	YONTVIEGRVILREKAKLVSNLSSORGSLYMEAGSTLDFVTPOPPQOPPANOLJTLS	564	QY 505 YONTVIEGRVILREKAKLVSNLSSORGSLYMEAGSTLDFVTPOPPQOPPANOLJTLS 564
Db	481	yrqvtiegrivirekalsvnsisqsgslymeagsl.dftvtpqppqpanqitls	540	Db 481 yrqvtiegrivirekalsvnsisqsgslymeagsl.dftvtpqppqpanqitls 540
QY	565	NLHLSLSLLANNAVTNPNTNPAQDHSHPAVGTSNTAGSVTISGPIFEDLDTAYDRD	624	QY 565 NLHLSLSLLANNAVTNPNTNPAQDHSHPAVGTSNTAGSVTISGPIFEDLDTAYDRD 624
Db	541	nhhsllsllannavtppptppqadqspqavqstagsvtisqsgslymeagsl	600	Db 541 nhhsllsllannavtppptppqadqspqavqstagsvtisqsgslymeagsl 600
QY	625	WLGSNQKTNVNLKQLGKPPNAPSDLTGNEMPKQYGSWSKLMADPNTANNGPYTLKA	684	QY 625 WLGSNQKTNVNLKQLGKPPNAPSDLTGNEMPKQYGSWSKLMADPNTANNGPYTLKA 684
Db	601	wlgsnqktnvnlkqlgkppnapsdtlgnempkqyggswkmladpntanngpytlka	660	Db 601 wlgsnqktnvnlkqlgkppnapsdtlgnempkqyggswkmladpntanngpytlka 660
QY	685	TWTKTGTPNPGPERVASLYPNSLWGSILDIRSAHSAIQASVDRGSRSYCRLWYSGVSNFYH	744	QY 685 TWTKTGTPNPGPERVASLYPNSLWGSILDIRSAHSAIQASVDRGSRSYCRLWYSGVSNFYH 744
Db	661	twtktgtpnpgpervaslypnslsldirsahsaiqasvdrgrscrglwvgsnifyh	720	Db 661 twtktgtpnpgpervaslypnslsldirsahsaiqasvdrgrscrglwvgsnifyh 720
QY	745	DRDALOGCYRYISGGYSUGANSYFGSSMFLGALATEVFRGRSKDYVVCRSNHHACIGSYLS	804	QY 745 DRDALOGCYRYISGGYSUGANSYFGSSMFLGALATEVFRGRSKDYVVCRSNHHACIGSYLS 804
Db	721	drdalqgyryisggysgansfgrsmflgatlevfrsksdvvcrsnhhacigsylys	780	Db 721 drdalqgyryisggysgansfgrsmflgatlevfrsksdvvcrsnhhacigsylys 780
QY	805	SKLYLNELRPFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSVPVKFDRCSSTHPK	924	QY 805 SKLYLNELRPFVQAEFSYADHESFTEEGDQARAFKSGHLLNLSVPVKFDRCSSTHPK 924
Db	841	sklylnelrpfqafefsadftegdqafrakshblnlsvpgkfdccsthpnk	900	Db 841 sklylnelrpfqafefsadftegdqafrakshblnlsvpgkfdccsthpnk 900
QY	925	YSPMAAYICDAYRTISGTTLJSHOEWTTDAFHLLRHGVVVRGSMYASLTSNIEVYGH	984	QY 925 YSPMAAYICDAYRTISGTTLJSHOEWTTDAFHLLRHGVVVRGSMYASLTSNIEVYGH 984
Db	901	ysmaayicdayrtisqettishqwtctafhllrhgvvvrsgnyasltsnievygh	960	Db 901 ysmaayicdayrtisqettishqwtctafhllrhgvvvrsgnyasltsnievygh 960
QY	985	GRYEYRDAASRGYGLSAGSKYRF	1006	QY 985 GRYEYRDAASRGYGLSAGSKYRF 1006
Db	961	gryeyrdaasrgyglsgsksvxf	982	Db 961 gryeyrdaasrgyglsgsksvxf 982
Sequence	982 AA.			

The present sequence is provided in a specification relating to compounds and methods for the treatment and diagnosis of chlamydia infection. The compounds provided include polypeptides and fusion proteins comprising immunogenic portions of Chlamydia antigens and DNA sequences encoding such polypeptides. They are useful for vaccinating against chlamydial infection, which causes pelvic inflammatory disease, trachoma, acute respiratory tract infections, atherosclerosis and heart disease.

Sequence 982 AA;

Qy	265 SFSRNTAVERFDGNVARYGGGTSYGNVAFNLNGKTILFLNNVASYAPVYIAKOPTSGOASNT	324
Qy	271 sfsrntaferdgnvarygggtsygnvafnlngktiflnnasyapvyaakoptsgasnt	330
Db	271 sfsrntaferdgnvarygggtsygnvafnlngktiflnnasyapvyaakoptsgasnt	330
Qy	275 SNNYDGGAFCKNAQAGSNNSGSFSDGCVWPFSSNVAGKGAIIYAKKLSVANCGP	384
Db	275 SNNYDGGAFCKNAQAGSNNSGSFSDGCVWPFSSNVAGKGAIIYAKKLSVANCGP	384
Qy	281 SNNYDGGAFCKNGAAGQGSNSGSVSFDGEVVFSNSVAAKGAIYAKKLSVANCGP	390
Db	281 SNNYDGGAFCKNGAAGQGSNSGSVSFDGEVVFSNSVAAKGAIYAKKLSVANCGP	390
Qy	285 VQFLRNIAANDGAYLGESEGLSLASADYGD1FDGDNLKRTRAKENAAADYNGTVSQAISIM	444
Db	285 VQFLRNIAANDGAYLGESEGLSLASADYGD1FDGDNLKRTRAKENAAADYNGTVSQAISIM	444
Db	291 vqflrnianggaiygelestslsdygdifdglnkrktkenaaadngtvvsqasim	450
Qy	295 GSGGKITTLLRAKAGHOIIFNDPIEMANGNNOPAQS SKLLKINDGE GTGDIVFANGSSTL	504
Db	295 GSGGKITTLLRAKAGHOIIFNDPIEMANGNNOPAQS SKLLKINDGE GTGDIVFANGSSTL	504
Db	301 vqflrnianggaiygelestslsdygdifdglnkrktkenaaadngtvvsqasim	504
Qy	305 YQDNVIEQGRIVRLRRAKLSVNSLSQTGGSLYMEASSTLDIVTPOPPQQPANQNLITLS	564
Db	305 YQDNVIEQGRIVRLRRAKLSVNSLSQTGGSLYMEASSTLDIVTPOPPQQPANQNLITLS	564
Db	311 yqnvrtieqgrivrlrreaklsvnslsqtgglymeastwdtvtppqqpaanqnlitls	570
Qy	325 NLHLSSLLIANNAVYNPPTNPPAQDSSHAPATGSTTAGSVTISGPPIFFEDLDTDAYDRD	624
Db	325 NLHLSSLLIANNAVYNPPTNPPAQDSSHAPATGSTTAGSVTISGPPIFFEDLDTDAYDRD	624
Db	331 nlhlslliannavnpptrippaqsdpshapav9stcagsvtisgpplifedldtdayrd	630
Qy	345 WLGSNOKINVILKQLQGTKPPANAPPSPDTLGNEMPKYGQGSNKLANDPNTANNGPYTLKA	684
Db	345 WLGSNOKINVILKQLQGTKPPANAPPSPDTLGNEMPKYGQGSNKLANDPNTANNGPYTLKA	684
Db	351 wlgsaqkinviklqlqgtkppanapsaltnempkqyqgsnklawoptantaqpytlka	690
Qy	365 TWTKGYNPGERVSLVPLVPLWGLSILDIDRAHSAQASVDRSICRGLWGSVSNFYH	744
Db	365 TWTKGYNPGERVSLVPLVPLWGLSILDIDRAHSAQASVDRSICRGLWGSVSNFYH	744
Db	371 tlwtkgynpgervslvplvplwgsildidrahsaqlqasvgrsycrg1lwsgvsnfyh	750
Qy	375 DRDAGQGYYRTSGGYSLGANSYFGSSMFLGIAFTTEYGRSKDYYVCRSNHHACIGSYLS	804
Db	375 DRDAGQGYYRTSGGYSLGANSYFGSSMFLGIAFTTEYGRSKDYYVCRSNHHACIGSYLS	804
Db	381 tlwtkgynpgervslvplvplwgsildidrahsaqlqasvgrsycrg1lwsgvsnfyh	810
Qy	385 TQQALCGSYLFGDAFIRASTGFGNQNMKTSYTFAAEESDVRDNNNCLAGEIGAGLPIVTP	864
Db	385 TQQALCGSYLFGDAFIRASTGFGNQNMKTSYTFAAEESDVRDNNNCLAGEIGAGLPIVTP	864
Db	391 tlwtkgynpgervslvplvplwgsildidrahsaqlqasvgrsycrg1lwsgvsnfyh	870
Qy	395 SKLYNELNLRPFVQAFSYADESFTPEQDQARAFKSHLNLNSVPGWKFDRCSSTHNPK	924
Db	395 SKLYNELNLRPFVQAFSYADESFTPEQDQARAFKSHLNLNSVPGWKFDRCSSTHNPK	924
Db	401 tlwtkgynpgervslvplvplwgsildidrahsaqlqasvgrsycrg1lwsgvsnfyh	930
Qy	405 YSFMAAYICDAYRTISGTETTLLSHOTWTWTDIAFHARHGYYVRGSNMYASLTNSIEYGH	984
Db	405 YSFMAAYICDAYRTISGTETTLLSHOTWTWTDIAFHARHGYYVRGSNMYASLTNSIEYGH	984
Db	411 tlwtkgynpgervslvplvplwgsildidrahsaqlqasvgrsycrg1lwsgvsnfyh	990
Qy	415 GRYEYRDASRGYGLSAGSKYRF	1006
Db	415 GRYEYRDASRGYGLSAGSKYRF	1006
Qy	421 991 gryeyrdasrgyglagsrfrf	1012
Sequence	421 1012 AA;	
Query Match	97.2%	Score 5090; DB 20; Length 1012;
Best Local Similarity	99.7%	Pred. No. 0;
Matches	979; Conservative	2; Mismatches 1; Indels 0; Gaps 0;
Job time:	1117 sec	
Search completed:	September 11, 2002, 12:32:05	

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 11, 2002, 12:28:33 ; Search time 25.91 Seconds

(without alignments) (730.832 Million cell updates/sec)

Title: US-09-841-132-190

Perfect score: 5238

Sequence: 1 MASWTCQQMGRDSSLVPHH.....YEYRDAASRGYGLSAGSKVRF 1006

Scoring table: BLOSUM62

Gapext 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 95%  
Maximum Match 100%  
Listing first 45 summaries

Database : P71R

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	4986.5	95.2	1013	2 G71460
				probable outer mem

#### ALIGNMENTS

RESULT	1
G71460	probable outer membrane protein G - Chlamydia trachomatis (serotype D, strain UW3/CX) C:Species: Chlamydia trachomatis C:accession: G71460 C:date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999 R; Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998 A:title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis A:reference number: A71570; MUID:9900809 A:accession: G71460 A:status: preliminary A:molecule type: DNA A:residues: 1-1013 <ARN> A:cross-references: GB:AE001360; GB:AF001273; NID:93329342; PIDN:AAC68469.1; PID:9332934 A:experimental source: serotype D, strain UW-3/CX C:Genetics: A:Gene: ppmg

QY	25 MIPQGIDGETLTVSPPTVIGDPSGTTVSAEGLJTLKNDNSTAALPLSCFGNLLGIFT 84
Db	31 MIPQGIDGETLTVSPPTVIGDPSGTTVSAEGLJTLKNDNSTAALPLSCFGNLLGIFT 90
QY	85 VLGRGHSLTENTIRTSNTNGAALSNSADGLIFTGKFELSFNSNQNSLLAVLPAAINKS 144
Db	91 VLGRGHSLTENIRTSNTNGAALSNSADGLIFTGKFELSFNSNQNSLLAVLPAAINKS 150
QY	145 QTPTTSTPENGTYSKTDILLNNEKFSYSNLVSGDGAIDAKSLTVOGISKLCVRF 204
Db	151 QTPTTSTPENGTYSKTDILLNNEKFSYSNLVSGDGAIDAKSLTVOGISKLCVRF 210
QY	205 NTAQADGGAQWVTSFAMANEAPTAFAVANVAGVRRGGTAAVQDGQOGQSSSTSTEDEPV 264
Db	211 NTAQADGGAQWVTSFAMANEAPTAFAVANVAGVRRGGTAAVQDGQOGQSSSTSTEDEPV 270
QY	265 SFSRNTAVERDGNTARYVGGGIYSYGNVAFNLNGTFLFLNNVASPVYIAKOPTSGQASNT 324
Db	271 SFSRNTAVERDGNTARYVGGGIYSYGNVAFNLNGTFLFLNNVASPVYIAKOPTSGQASNT 330
QY	325 SNNTGGGAIFCKNGAQ-AGSNNSGSYSDEGEVGFSSVNAAGKGATYAKKSYANG 383
Db	331 SDNTGGGAIFCKNGAQ-AGSNNSGSYSDEGEVGFSSVNAAGKGATYAKKSYANG 390
QY	384 PVOFLRNLANDGGAYLGESELISLADYGDIFDGNLKRKTAKENAADINGVUTSSQAT 443
Db	391 PVOFLRNLANDGGAYLGESELISLADYGDIFDGNLKRKTAKENAADINGVUTSSQAT 450
QY	444 MGSGKITTIRAKAGHQILFNDPTEMANENNQPOQSKLKKINGEGYGTIDYFVANGGST 503
Db	451 MGSGKITTIRAKAGHQILFNDPTEMANENNQPOQSKLKKINGEGYGTIDYFVANGGST 510
QY	504 LYQNTYIEGRIVREKAKLVSNSLQTGGSLYMEAGSPLDFVYPOPOQPRPAQNLTL 563
Db	511 LYQNTYIEGRIVREKAKLVSNSLQTGGSLYMEAGSPLDFVYPOPOQPRPAQNLTL 570
QY	564 SNLHLSLSSLLANNAWNTNPNTPAQDOSHAPVAGSTAGSVTISGPPIFEDDDTAYDR 623
Db	571 SNLHLSLSSLLANNAWNTNPNTPAQDOSHAPVAGSTAGSVTISGPPIFEDDDTAYDR 630
QY	624 DWLGSNQKINVKLQLQGKTPKANAPASDPLTGLNEMPKYGVGSWNLAWPDNTANNGPYRKL 683
Db	631 DWLGSNQKINVKLQLQGKTPKANAPASDPLTGLNEMPKYGVGSWNLAWPDNTANNGPYRKL 690
QY	684 ATWTKTGYRGPVERVAVLPSLNSLDIRSRAHSIAQASYDERSYCRGLMWVGSVNFY 743
Db	691 ATWTKTGYRGPVERVAVLPSLNSLDIRSRAHSIAQASYDERSYCRGLMWVGSVNFY 750
QY	744 HDRADLGQGRYITSGGSYGANSYFGSSMGLAFTEVGRSKDWWVCRSNHACIGSYL 803
Db	751 HDRADLGQGRYITSGGSYGANSYFGSSMGLAFTEVGRSKDWWVCRSNHACIGSYL 810
QY	804 STQOALCGSYLFGDAFTRASYFGNQHMTSYFAEESDVRWDNCLAGETGAGLPLVIT 863
Db	811 STQOALCGSYLFGDAFTRASYFGNQHMTSYFAEESDVRWDNCLAGETGAGLPLVIT 870
QY	864 PSKLYLNELRPFYQAEFADHSEFTESDQAFRAFKSCHLNLNSVPGVCFEDCSSTPHN 923
Db	871 PSKLYLNELRPFYQAEFADHSEFTESDQAFRAFKSCHLNLNSVPGVCFEDCSSTPHN 930
QY	924 KYSMAYICDAYTISGETTSLSHQETWTDAFHARHGVVURGSMYASLTSNIEVY 983
Db	931 KYSMAYICDAYTISGETTSLSHQETWTDAFHARHGVVURGSMYASLTSNIEVY 990
QY	984 HGRVEYRDAASRGYGLSAGSKVRF 1006
Db	991 HGRVEYRDAASRGYGLSAGSKVRF 1013

Query Match 95.2%; Score 4986.5; DB 2; Length 1013;  
Best Local Similarity 97.4%; Pred. No. 2.3e-294;  
Matches 957; Conservative 15; Mismatches 10; Indels 1; Gaps 1;

Search completed: September 11, 2002, 12:32:59  
Job time: 266 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 11, 2002, 12:32:28 ; search time 15.1 Seconds

(without alignments)  
2579.594 Million cell updates/sec

Title: US-09-841-132-190

Perfect score: 5238

Sequence: 1 MASMTGGQOMGRDSSLVPHH.....YEYRDASRGYGLSAGSKYRF 1006

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 95%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No. Score Query No. Description

Match Length 1013

DB ID 1

ID PMPG\_CHLTR

STANDARD; PRT; 1013 AA.

ID 084879;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Probable outer membrane protein PmpG precursor (Polymorphic membrane protein G).

GN PMPG OR Cr871.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

ID NCBI\_TaxID=813;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=D/IN-3/CX;

RX MEDLINE=9900809; PubMed=9784136;

RA Stephens R.S., Kaiman S., Lammel C.J., Fan J., Marathe R., Aravind L.,

RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,

RA Davis R.W.;

RT "Genome sequence of an obligate intracellular pathogen of humans:

RT Chlamydia trachomatis";

RL science 282:754-759(1998);

CC -1- SUBCELLULAR LOCATION: CELL WALL SURFACE (ELEMENTARY BODIES)

CC (POTENTIAL).

CC BELONGS TO THE PMP OUTER MEMBRANE PROTEIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Bioinformatics Institute and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AEG01360; AAC68469; 1;

CC DR InterPro; IPR003368; DUF145.

CC DR InterPro; IPR003357; OMP.

CC DR Pfam; PF02415; DUF45; 1.

CC DR Pfam; PF02385; OMP\_1;

CC KW Outer membrane; Signal; Multigene family; Complete proteome.

CC FT SIGNAL 1 27 POTENTIAL.

CC FT CHAIN 28 1013 PROBABLE OUTER MEMBRANE PROTEIN PMPG.

CC SQ SEQUENCE 1013 AA; 107366 MW; F0927743CO651DD CRC64;

Query Match 95.2%; Score 4986.5; DB 1; Length 1013;

Best Local Similarity 97.4%; Pred. No. 4.7e-286; Indels 1; Gaps 1;

Matches 957; Conservative 15; Mismatches 10; Indels 1; Gaps 1;

Qy 25 MIPQIYDGGETLTYSFPYTVIGDPSGTVFSAGELTAKNLDNSAALPLSCFGNLLGSFT 84

Db 31 MIPQIYDGGETLTYSFPYTVIGDPSGTVFSAGELTAKNLDNSAALPLSCFGNLLGSFT 90

Qy 85 VLGRGHSLTFENITISTNGAALSNSAAGDLETTIGFKELSFSNSNNSLAVLPATTNGS 144

Db 91 VLGRGHSLTFENITISTNGAALSNSAAGDLETTIGFKELSFSNSNNSLAVLPATTNGS 150

Qy 145 QTPPTTSPSNGTLYSKTDLNNEKEFSYNSLVSQDGDAIDAKSLSLTVOGISKLCVFQE 204

Db 151 QTPPTTSPSNGTLYSKTDLNNEKEFSYNSLVSQDGDAIDQKSLSLVQGISKLCVFQE 210

Qy 205 NTADGGACQVWTSFSAMANEAFIAFYANVAGYRGGGTAAVQDQGQYSSSSTEDPVV 264

Db 211 NTADGGACQVWTSFSAMANEAFIAFYANVAGYRGGGTAAVQDQGQYSSSSTEDPVV 270

Qy 265 SFSTNTAVEFDGNATARYGGIYSTGNVAPFLNGTFLFLNVASVYIAAKQPSQASIT 324

Db 271 SFSTNTAVEFDGNATARYGGIYSTGNVAPFLNGTFLFLNVASVYIAAEQPTNGQASIT 330

Qy 325 SNNYGDGGAIFCKNGAQ-AGSNNSGSVSDGEVVFESSNVAAKGGAIAYAKLISVANG 383

Db 331 SDNYGDGGAIFCKNGAQAAQGSNNNSGSVSPDGEVVFESSNVAAKGGAIAYAKLISVANG 390

Qy 384 PVQFLRNIAANDGGATYLGESGELSISADYGDIIIFDGNLKRTAKENAADVNGVTSSQASL 443

Db 391 PVQFLGNIAANDGGATYLGESGELSISADYGDIIIFDGNLKRTAKENAADVNGVTSSQASL 450

Qy 444 MGSGSKITTLRAKGHQLFLNDPEMANNNOPQSSKLLINGEGITGDIVPANGSST 503

Db 451 MGSGSKITTLRAKGHQLFLNDPEMANNNOPQSSKLLINGEGITGDIVPANGSST 510

Qy 504 LYQNTVIEGRIVLREKAKLVSNSLSTQGGSILYMEPLDFVTPQQPQQPAPANOLITL 563

Db 511 LYQNTVIEGRIVLREKAKLVSNSLSTQGGSILYMEAGSTLDFVTPQQPQQPAPANOLITL 570

Qy 564 SNLHSLSLLNANAVNTPNPAODSHPAVIGSTTAGSVTISGPFFBDLDTAYDRY 623

Db 571 SNLHSLSLLNANAVNTPNPAQDSHPAVIGSTTAGSVTISGPFFBDLDTAYDRY 630

Qy 624 DWLGSNOKINVKLQLGTPPANPSDTLGNENPKYQYQGSNWLAWDNTANGPYTUK 683

Db 631 DWLGSNOKDVLKQLGTPSANSPSDTLGNENPKYQYQGSNWLAWDNTANGPYTUK 690

Qy 684 ATWTKTGYNPGPVERASVLPNSLMSILDIRSNSAIAQSVDGRSYCRGLWSSVSNFFY 743

Db 691 ATWTKTGYNPGPVERASVLPNSLMSILDIRSNSAIAQSVDGRSYCRGLWSSVSNFFY 750

Qy 744 HDRDALGQGYRYISGGYSGLANSYFGSSMFLGLAPTEVGRSKDVFVCRSNHHAGIGSYL 803

GenCore version 4.5  
 Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on:

September 11, 2002, 12:32:08 ; Search time 40.85 Seconds

(without alignments)  
 4259.249 Million cell updates/sec

Title: US-09-841-132-190  
 Perfect score: 5238  
 Sequence: 1 MASMTGQQMGRDSLVPHH..... YEYRDAASRGYGLSAGSKVRF 1006

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 95%  
 Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_19:  
 1: sp\_archea:  
 2: sp\_bacteria:  
 3: sp\_fungi:  
 4: sp\_human:  
 5: sp\_invertebrate:  
 6: sp\_mammal:  
 7: sp\_mic:  
 8: sp\_organelle:  
 9: sp\_phage:  
 10: sp\_plant:  
 11: sp Rodent:  
 12: sp\_virus:  
 13: sp\_vertebrate:  
 14: sp\_unclassified:  
 15: sp\_virus:  
 16: sp\_bacteriap:  
 17: sp\_archeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
------------	-------------	--------------	-------	-------------

No matches found

Search completed: September 11, 2002, 12:36:59  
 Job time: 291 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 11, 2002, 12:25:13 ; Search time 17.9 Seconds

(without alignments)  
1372.746 Million cell updates/sec

Title: US-09-841-132-190  
Perfect score: 5238  
Sequence: 1 MASMTGGQQMGRDSSLVPHH.....YEYRDASRGYGLSAGSKVRF 1006

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post processing: Minimum Match 95%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:/\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:/\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:/\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:/\*  
5: /cgn2\_6/ptodata/2/1aa/PC7US\_COMB.pep:/\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:/\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match Length	DB ID	Description
					-----

No matches found

Search completed: September 11, 2002, 12:32:25  
Job time: 432 sec

